

# Estimation of Phase Noise in Oscillators with Colored Noise Sources

M. Reza Khanzadi, *Student Member, IEEE*, Rajet Krishnan, *Student Member, IEEE*, and Thomas Eriksson

**Abstract**—In this letter we study the design of algorithms for estimation of phase noise (PN) with colored noise sources. A soft-input maximum a posteriori PN estimator and a modified soft-input extended Kalman smoother are proposed. The performance of the proposed algorithms are compared against those studied in the literature, in terms of mean square error of PN estimation, and symbol error rate of the considered communication system. The comparisons show that considerable performance gains can be achieved by designing estimators that employ correct knowledge of the PN statistics.

## I. INTRODUCTION

OSCILLATOR PHASE NOISE (PN) results in challenging synchronization issues which degrade the performance of communication systems [1], [2]. Demands for high data rates motivate the use of high-order modulation schemes in such systems. Nevertheless, PN severely limits the performance of systems that employ dense constellations.

The problem of PN estimation has been widely studied during the last decades (see [1], [3], [4] and the references therein). In [4], a feedforward PN estimation-symbol detection algorithm is presented, while iterative methods for joint phase estimation and symbol detection are studied in [1].

In prior studies, PN is modeled as a discrete random walk with uncorrelated (white) Gaussian increments between each time instant (i.e., the discrete Wiener process). This model results from using oscillators with white noise sources [5]. However, numerous studies show that real oscillators also contain colored noise sources, and PN is accurately modeled as a random walk with correlated (colored) Gaussian increments [5]–[7].

In this letter, we propose techniques to estimate PN from real oscillators with white and colored noise sources, in a single antenna-single carrier communication system. We first derive a general soft-input maximum a posteriori (MAP) PN estimator that is optimal in terms of the mean square error (MSE). Then, a modified soft-input extended Kalman smoother is proposed that can be used for estimation of PN with colored increments. The proposed Kalman smoother is observed to perform close to the MAP estimator in several interesting scenarios, with a significantly reduced complexity. Further, we compare the proposed methods with state of the art techniques. The proposed estimators jointly estimate the PN samples of a block of received signals, which improves the estimation performance compared to sequential PN estimation algorithms previously studied (e.g., [3]). Our estimators can be used in feedforward or iterative designs for the estimation of PN with white and colored increments.

The authors are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden. M. Reza Khanzadi is also with the Department of Microtechnology and Nanoscience, Chalmers University of Technology. Email: {khanzadi, rajet, thomase}@chalmers.se.

## II. SYSTEM MODEL

Consider the transmission of a block of  $K$  data symbols over an additive white Gaussian noise (AWGN) channel, affected by random PN. The channel coefficient from the transmitter to receiver antenna is assumed to be constant over the transmitted block, and it is estimated and compensated by employing a known training sequence that is transmitted prior to the data symbols [2]. In the case of perfect timing synchronization, the received signal after sampling the output of matched filter can be modeled as in [1]

$$y_k = s_k e^{j\theta_k} + w_k, \quad k \in \{1, \dots, K\}, \quad (1)$$

where  $\theta_k$  represents the PN affecting the  $k$ th received signal due to noisy transmitter and receiver local oscillators, and  $w_k$  is a realization of the independent and identically distributed (i.i.d.) zero-mean complex circularly symmetric AWGN with variance  $\sigma^2$ . In this model,  $\mathbf{y} = \{y_k\}_{k=1}^K$  is the sequence of received signals and  $\mathbf{s} = \{s_k\}_{k=1}^K$  is the transmitted symbol sequence divided in two sets of symbols,  $K_p$  known pilot symbols, and  $K - K_p$  unknown data symbols. We model the pilot and data symbols in general as

$$s_k = \hat{s}_k + \epsilon_k, \quad (2)$$

where  $\hat{s}_k$  is the soft detected symbol and  $\epsilon_k$  models the uncertainty of  $s_k$  as an i.i.d zero-mean circularly symmetric AWGN with variance  $\sigma_{\epsilon_k}^2$ . Such a modeling choice is commonly used in the literature [8]. For the pilot symbols,  $\sigma_{\epsilon_k}^2 = 0$  since they are known. Using (1) and (2), the received signal can be rewritten as

$$y_k = \hat{s}_k e^{j\theta_k} + \underbrace{\epsilon_k e^{j\theta_k}}_{\triangleq \tilde{w}_k} + w_k, \quad k \in \{1, \dots, K\}, \quad (3)$$

where  $\tilde{w}_k$  is the new observation noise. As  $\epsilon_k$  is modeled circularly symmetric,  $\tilde{w}_k \sim \mathcal{CN}(0, \sigma_k^2 \triangleq \sigma^2 + \sigma_{\epsilon_k}^2)$ .

The PN samples are modeled by a random-walk as

$$\theta_k = \theta_{k-1} + \zeta_{k-1}, \quad (4)$$

where the phase increment process  $\zeta_k$  is a zero-mean Gaussian random process. Recent studies of the PN in oscillators with colored noise sources show that the PN increments can be correlated over time [5]–[7]. Hence, we consider a general case where the autocorrelation function of  $\zeta_k$ , denoted as  $R_\zeta(l)$ , is known a priori. Note that the Wiener PN model extensively used in the literature is a special case of the proposed model, with uncorrelated (white) phase increments [5].

*Notations:* Italic letters ( $x$ ) are scalar variables, boldface letters ( $\mathbf{x}$ ) are vectors, uppercase boldface letters ( $\mathbf{X}$ ) are matrices,  $[\mathbf{X}]_{a,b}$  denotes the  $(a, b)^{th}$  entry of matrix  $\mathbf{X}$ ,  $\text{diag}(\mathbf{X})$  denotes the diagonal elements of matrix  $\mathbf{X}$ ,  $\mathbb{E}[\cdot]$  denotes the statistical expectation operation,  $\mathcal{CN}(x; \mu, \sigma^2)$  denotes the complex proper Gaussian distribution with variable  $x$ , mean  $\mu$ , and variance  $\sigma^2$ ,  $\log(\cdot)$  denotes the natural logarithm,  $\Re\{\cdot\}$ ,  $\Im\{\cdot\}$ , and  $\arg\{\cdot\}$  are the real part, imaginary part, and angle of complex-valued numbers, and  $(\cdot)^*$  and  $(\cdot)^T$  denote the conjugate and transpose, respectively.

### III. PHASE NOISE ESTIMATION

In the sequel, we propose two methods for joint estimation of  $K$ -dimensional PN vector  $\boldsymbol{\theta} = \{\theta_k\}_{k=1}^K$ , that further would be used for data detection. First, we derive a MAP estimator. Thereafter, we propose an approach for modification of (4), such that smoothing algorithms (e.g., Kalman smoother) with a lower complexity than MAP can be used for estimation.

#### A. Proposed MAP Estimator

Let  $f(\boldsymbol{\theta}|\mathbf{y})$  denote the a posteriori distribution of PN vector  $\boldsymbol{\theta}$ , given the observation vector  $\mathbf{y}$ . The MAP estimator of  $\boldsymbol{\theta}$  is determined as

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \log(f(\boldsymbol{\theta}|\mathbf{y})) = \arg \max_{\boldsymbol{\theta}} \log(f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})), \quad (5)$$

where we define  $\ell(\boldsymbol{\theta}) \triangleq \log(f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta}))$ . To solve this optimization, we first need to find the likelihood,  $f(\mathbf{y}|\boldsymbol{\theta})$ , and prior distribution of  $\boldsymbol{\theta}$ ,  $f(\boldsymbol{\theta})$ . As both  $w_k$  and  $\epsilon_k$  are i.i.d, and  $y_k$  only depends on  $\theta_k$  according to (3), the likelihood function can be written as

$$f(\mathbf{y}|\boldsymbol{\theta}) = \prod_{k=1}^K f(y_k|\theta_k) = \prod_{k=1}^K f(y_k|\theta_k), \quad (6)$$

where

$$\begin{aligned} f(y_k|\theta_k) &= \mathcal{CN}(y_k; \hat{s}_k e^{j\theta_k}, \sigma_k^2) \\ &= \frac{1}{\sigma_k^2 \pi} \exp\left(-\frac{|y_k - \hat{s}_k e^{j\theta_k}|^2}{\sigma_k^2}\right). \end{aligned} \quad (7)$$

In order to find the prior distribution  $f(\boldsymbol{\theta})$  of the PN vector, we use the random walk model in (4) that results in a general PN incremental form of

$$\theta_k = \theta_1 + \sum_{i=1}^{k-1} \zeta_i, \quad (8)$$

where  $\theta_1$  (PN of the first symbol in the block) is modeled as a zero-mean Gaussian random variable with a high variance<sup>1</sup>, denoted as  $\sigma_{\theta_1}^2$ . Based on (8), we can show that  $\boldsymbol{\theta}$  has a multivariate Gaussian distribution  $f(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta}; \mathbf{0}, \mathbf{C})$  where elements of the covariance matrix  $\mathbf{C}$  can be computed as

$$\begin{aligned} [\mathbf{C}]_{m,m'} &= \mathbb{E}[(\theta_m - \mathbb{E}[\theta_m])(\theta_{m'} - \mathbb{E}[\theta_{m'}])] \\ &= \sigma_{\theta_1}^2 + \sum_{l=1}^{m-1} \sum_{l'=1}^{m'-1} R_\zeta(l-l'), \quad m, m' \in \{1, \dots, K\}. \end{aligned} \quad (9)$$

From (7) and the multivariate Gaussian prior of  $\boldsymbol{\theta}$  we obtain

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^K \frac{2}{\sigma_k^2} \Re\{y_k \hat{s}_k^* e^{-j\theta_k}\} - \frac{1}{2}(\boldsymbol{\theta}^T \mathbf{C}^{-1} \boldsymbol{\theta}) + \text{const.} \quad (10)$$

To find the maximizer of (10), an exhaustive grid-search over all possible values of  $\boldsymbol{\theta}$  can be used. However, the complexity of this method increases exponentially with the length of  $\boldsymbol{\theta}$ . The stationary point of this optimization can analytically be found as the root of the gradient of  $\ell(\boldsymbol{\theta})$  with respect to  $\boldsymbol{\theta}$ ,

$$\mathbf{g}(\boldsymbol{\theta}) \triangleq 2 \left[ \left\{ \frac{\Im\{y_k \hat{s}_k^* e^{-j\theta_k}\}}{\sigma_k^2} \right\}_{k=1}^K \right]^T - \mathbf{C}^{-1} \boldsymbol{\theta} = \mathbf{0}. \quad (11)$$

<sup>1</sup>We consider a flat non-informative prior [9] for the initial PN value, modeled by a Gaussian distribution with a high variance that behaves similar to a flat prior over a certain interval.

In order to solve  $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ , which is a non-linear system of equations, we use the Newton-Raphson method whose  $n$ th iteration is given by

$$\hat{\boldsymbol{\theta}}^{(n+1)} = \hat{\boldsymbol{\theta}}^{(n)} - \mathbf{H}^{-1}(\hat{\boldsymbol{\theta}}^{(n)}) \mathbf{g}(\hat{\boldsymbol{\theta}}^{(n)}), \quad (12)$$

where  $\mathbf{H}(\boldsymbol{\theta})$  denotes the Hessian matrix,

$$\mathbf{H}(\boldsymbol{\theta}) \triangleq - \left( 2 \text{diag} \left( \left[ \left\{ \frac{\Re\{y_k \hat{s}_k^* e^{-j\theta_k}\}}{\sigma_k^2} \right\}_{k=1}^K \right] \right) + \mathbf{C}^{-1} \right). \quad (13)$$

We iterate till an accurate value of the root is reached.

To show that this algorithm reaches a global maximum, we first prove that  $\ell(\boldsymbol{\theta})$  is a concave function in moderate and high signal to noise ratio (SNR) regimes. In (13),  $\mathbf{C}^{-1}$  is the inverse of a covariance matrix, thus it is a positive-definite matrix. If the first term of the sum in (13) is also positive-definite (or positive-semidefinite), the Hessian becomes negative-definite, thus implying that  $\ell(\boldsymbol{\theta})$  is a concave function. For the first term of (13) to be positive-semidefinite,  $\Re\{y_k \hat{s}_k^* e^{-j\theta_k}\}$  must be greater than or equal to zero. Exploiting Eq. (3) we get

$$\Re\{y_k \hat{s}_k^* e^{-j\theta_k}\} = |\hat{s}_k|^2 + \Re\{\tilde{s}_k^* \tilde{w}_k\} \geq 0, \quad (14)$$

where  $\tilde{w}_k \triangleq e^{-j\theta_k} \tilde{w}_k$  with the same statistics as  $\tilde{w}_k$ . It is clear that for moderate and high SNR, (14) is satisfied with a high probability, and consequently  $\ell(\boldsymbol{\theta})$  is a concave function.

In the low-SNR regime,  $\ell(\boldsymbol{\theta})$  is not necessarily concave. Therefore, we propose an approach to initiate the iterations with a guess which is fairly close to the optimal point. This ensures that the method does not get trapped in a local maxima, far from the global maximum. Moreover, employing a good initial guess speeds up the convergence of the algorithm at any SNR. In this respect, we first find the maximum likelihood (ML) estimate of the PN samples for the pilot symbols. For any  $s_k$  in the pilot set, the ML estimator of the  $k$ th PN sample can be computed as  $\hat{\theta}_k^{\text{ML}} = \arg\{y_k s_k^*\}$ . Then, we form our initial estimate of the PN vector,  $\hat{\boldsymbol{\theta}}^{(0)}$ , as the linear interpolation of the ML estimated PN values.

The MAP estimator is an optimal minimum mean square estimator if its MSE attains the Bayesian Cramér-Rao bound (BCRB) [9]. In the Appendix, we analytically derive the MSE of the MAP, and show that it is approximately equal to the BCRB of the PN estimation. Our simulation results in Sec. IV also confirms this (Fig. 2).

Although, the proposed MAP estimator gives an optimal estimate of  $\boldsymbol{\theta}$ , as we can see in (12), it involves inversion of  $K \times K$  matrices that may raise some complexity issues. In the next section, we propose an approach to modify the PN model and reduce the complexity.

#### B. Auto Regressive Model of Colored Phase Noise Increments

In general, the PN increment process can be modeled with a  $p$ th-order auto regressive (AR) process as follows

$$\zeta_k = \sum_{i=1}^p \alpha_i \zeta_{k-i} + \Delta_k, \quad (15)$$

where  $\alpha_i$  are the coefficients of the AR model and  $\Delta_k$  is modeled as a zero-mean white noise process with variance  $\sigma_{\Delta}^2$ . For a given autocorrelation  $R_\zeta(l)$  and AR order  $p$ , the optimal  $\alpha_i$  and  $\sigma_{\Delta}^2$  can be computed using algorithms such as the Levinson-Durbin recursion. We modify the state equation (4) with the AR model (15), which results in an augmented state equation,

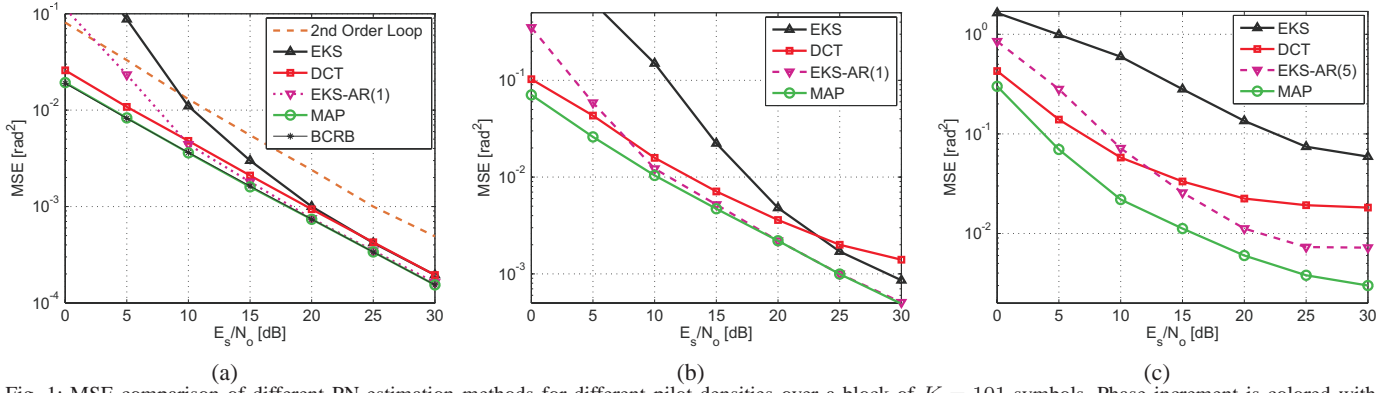


Fig. 1: MSE comparison of different PN estimation methods for different pilot densities over a block of  $K = 101$  symbols. Phase increment is colored with variance  $R_{\zeta}(0) = 10^{-3}[\text{rad}^2]$ . (a) Data-aided case, where all symbols are pilots. (b) Pilot density of 21%. (c) Pilot density of 6%.

$$\begin{bmatrix} \theta_k \\ \zeta_k \\ \zeta_{k-1} \\ \vdots \\ \zeta_{k-p} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ 0 & \alpha_1 & \alpha_2 & \dots & \alpha_p \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}}_{\triangleq \mathbf{F}} \underbrace{\begin{bmatrix} \theta_{k-1} \\ \zeta_{k-1} \\ \zeta_{k-2} \\ \vdots \\ \zeta_{k-p-1} \end{bmatrix}}_{\triangleq \mathbf{x}_{k-1}} + \underbrace{\begin{bmatrix} 0 \\ \Delta_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\triangleq \Delta_{k-1}}, \quad (16)$$

where  $\mathbf{x}_k$  is the new state vector that has a higher dimension compared to our original state variable  $\theta_k$ , and  $\Delta_k$  denotes the new process noise that is white, with covariance

$$\mathbb{E}[\Delta_k \Delta_k^T] = \text{diag}([0, \sigma_{\Delta}^2, 0, \dots, 0]). \quad (17)$$

We can use the new state equation (16) along with the observation model (3) to estimate  $\theta$  by Kalman filtering/smoothing [9]. For the colored PN increments with a long memory, normally a high-order AR model is needed that results in a high dimensional state equation (16). In order to reduce complexity, we approximate the colored PN increments with a low-order AR process. Numerical simulations in Sec. IV show that even with such an approximation, the modified extended Kalman smoother (EKS) perform close to the proposed MAP estimator, in several scenarios of interest.

#### IV. NUMERICAL RESULTS

We now study the performance of our proposed estimators and compare with that of those available in the literature (e.g., [1], [3], [4]). We consider transmission of a block of  $K = 101$  16-QAM modulated symbols, with uniformly distributed pilots. For simulation of the PN with colored noise increments, we use the results of [6], [7], where the autocorrelation function of the PN increments for oscillators with a colored noise source (flicker noise) has been derived. We set the parameters such that the variance of the PN increments becomes  $R_{\zeta}(l=0) = 10^{-3}[\text{rad}^2]$ .

It can be seen in Fig. 1-(a) that the MSE of the proposed MAP estimator reaches the BCRB [7] in the data-aided case, where all symbols are pilots. We stop the optimization algorithm when the gradient is sufficiently small (here  $|\mathbf{g}(\theta)| < 10^{-6}$ ). We observe that the number of required iterations for satisfying any level of accuracy depends on various parameters such as the block length, the PN statistics, the pilot density and the SNR. In general, for most practical scenarios less than 5 iterations suffice. For instance, for simulations of Fig. 1 with

6% pilot density, at SNRs 0 and 30 dB, on average 4.3 and 2.95 iterations are required, respectively.

The MSE of the PN estimator proposed in [4], based on interpolation of the PN estimates of the pilot symbols using discrete cosine transform (DCT), is close to the MAP estimator in the data-aided case. However, decreasing the pilot density in Fig. 1-(b) and (c) to 21% and 6%, shows that the DCT-based estimator performs poorly in low-pilot density scenarios. The reason is that this estimator is blind to the statistics of the PN process that limits its performance when the received signals are not reliable.

Using an EKS designed for white PN increments in the colored case results in large MSEs in low-SNR and low-pilot density cases. In high-pilot density scenarios, the observations are reliable and the EKS performs close to the MAP. When the pilot density is low, the tested EKS relies on PN statistics that are not matched to the real process, which results in large MSEs. Now consider the modified EKS designed with the low-order AR approximation. Using a first-order AR model, the modified EKS reaches the MSE of the MAP in the data-aided and 21% pilot density cases. With 6% pilot density, where the modified EKS relies more on the PN statistics, a higher order AR model is needed to improve the performance ( $p > 5$ ). Fig. 1-(a) also shows the data-aided MSE of the second-order phase tracking loop [3].

Fig. 2 shows that the simulated MSE of the MAP estimator reaches the BCRB. Fig. 3 and 4 compare the effect of using the discussed estimators on symbol error rate (SER) of the system, after three iterations between the PN estimators and a Euclidian-distance-based symbol detector. Mean and variance of the soft symbols are calculated as the mean and variance of the symbols' a posteriori probabilities. In both scenarios, the MAP estimator outperforms other estimators. It can also be seen that in the 6% density compared to 21% scenario, a higher order AR model is needed for more accurate approximation of the colored PN increments. In addition to the estimators in Fig. 1, we also study the performance of an iterative receiver that is designed based on the sum-product algorithm (SPA) [1]. This SPA-based receiver performs extremely well in presence of the Wiener PN, but it is not designed for PN with colored increments.

#### V. CONCLUSIONS

In this letter, we showed that deriving the soft-input maximum a posteriori (MAP) estimator for estimation of phase noise (PN) in oscillators with colored noise sources is a



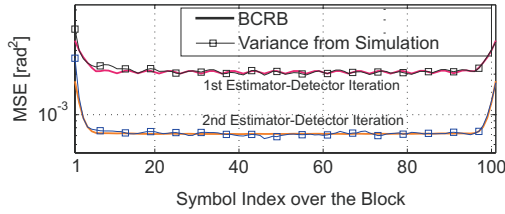


Fig. 2: Simulated variance of MAP vs. BCRB. Pilot density 21%,  $R_\zeta(0) = 10^{-3}[\text{rad}^2]$ , SNR = 20[dB].

concave optimization problem at moderate and high SNRs. Further, we showed that the modified soft-input extended Kalman smoother with low-order AR approximation of the colored PN increments performs close to the MAP in several scenarios. From simulations, we observed that considerable performance gain can be achieved by using the proposed estimators compared to estimators that lack correct statistic of the PN. The gain is more significant in low-SNR or low-pilot density scenarios.

#### APPENDIX

Here, we find the mean and covariance of the MAP estimation error, defined as  $\psi \triangleq (\theta^\dagger - \hat{\theta})$ , where  $\theta^\dagger$  denotes the true value of  $\theta$ . We first write the Taylor expansion of  $\mathbf{g}(\theta)$  around  $\theta^\dagger$  and evaluate it at  $\hat{\theta}$ . Assuming that  $\hat{\theta}$  is close to  $\theta^\dagger$ , we can neglect the higher order terms and obtain

$$\mathbf{g}(\hat{\theta}) \approx \mathbf{g}(\theta^\dagger) + \mathbf{H}(\theta^\dagger)(\hat{\theta} - \theta^\dagger). \quad (18)$$

Note that  $\hat{\theta}$  is the root of  $\mathbf{g}(\theta) = \mathbf{0}$ . Therefore,

$$\theta^\dagger = \hat{\theta} + \mathbf{H}^{-1}(\theta^\dagger)\mathbf{g}(\theta^\dagger), \quad (19)$$

where  $\psi \triangleq (\theta^\dagger - \hat{\theta}) = \mathbf{H}^{-1}(\theta^\dagger)\mathbf{g}(\theta^\dagger)$  is the estimation error term whose mean is calculated as

$$\mathbb{E}[\psi] = \mathbb{E}[\mathbf{H}^{-1}(\theta^\dagger)\mathbf{g}(\theta^\dagger)]. \quad (20)$$

Setting the value of  $y_k$  from (3) in (11) and (13) gives

$$\mathbf{g}(\theta^\dagger) = 2 \left[ \left\{ \frac{\Im\{\hat{s}_k^* \tilde{w}_k\}}{\sigma_k^2} \right\}_{k=1}^K \right]^T - \mathbf{C}^{-1}\theta^\dagger, \quad (21a)$$

$$\mathbf{H}(\theta^\dagger) = -2 \text{diag} \left( \left[ \left\{ \frac{|\hat{s}_k|^2 + \Re\{\hat{s}_k^* \tilde{w}_k\}}{\sigma_k^2} \right\}_{k=1}^K \right] \right) - \mathbf{C}^{-1}, \quad (21b)$$

where  $\tilde{w}_k \triangleq e^{-j\theta_k} \tilde{w}_k$  with the same statistics as  $\tilde{w}_k$ . It is clear that  $\mathbf{H}(\theta^\dagger)$  and  $\mathbf{g}(\theta^\dagger)$  are independent. Therefore,

$$\mathbb{E}[\psi] = \mathbb{E}[\mathbf{H}^{-1}(\theta^\dagger)]\mathbb{E}[\mathbf{g}(\theta^\dagger)] = \mathbf{0}, \quad (22)$$

where the second equality is true because  $\mathbb{E}[\mathbf{g}(\theta^\dagger)] = \mathbf{0}$ .

The covariance matrix of  $\psi$  is determined as

$$\Sigma = \mathbb{E}[\psi\psi^T] = \mathbb{E}[\mathbf{H}^{-1}(\theta^\dagger)\mathbf{g}(\theta^\dagger)\mathbf{g}^T(\theta^\dagger)\mathbf{H}^{-1}(\theta^\dagger)]. \quad (23)$$

In (21b), it is possible to neglect  $\Re\{\hat{s}_k^* \tilde{w}_k\}$  compared to  $|\hat{s}_k|^2$  for moderate and high SNRs. Therefore,  $\mathbf{H}(\theta^\dagger)$  is approximated as

$$\mathbf{H}(\theta^\dagger) \approx \tilde{\mathbf{H}}(\theta^\dagger) = -2 \text{diag} \left( \left[ \left\{ \frac{|\hat{s}_k|^2}{\sigma_k^2} \right\}_{k=1}^K \right] \right) - \mathbf{C}^{-1}, \quad (24)$$

which is a deterministic matrix. Thus, the expectation in (23) is only over  $\mathbf{g}(\theta^\dagger)\mathbf{g}^T(\theta^\dagger)$ . Based on (21a) and after straightforward mathematical manipulation,

$$\mathbb{E}[\mathbf{g}(\theta^\dagger)\mathbf{g}^T(\theta^\dagger)] = -\tilde{\mathbf{H}}(\theta^\dagger), \quad (25)$$

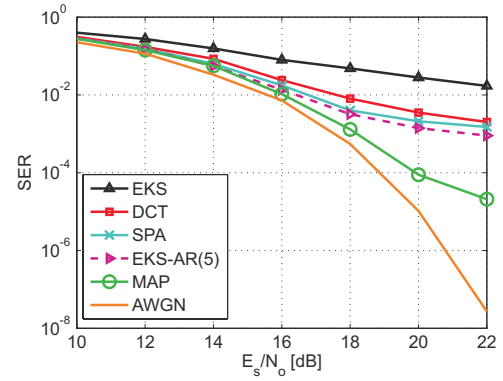


Fig. 3: SER comparison of different PN estimation methods after three estimation-detection iterations. Pilot density 6% and  $R_\zeta(0) = 10^{-3}[\text{rad}^2]$ .

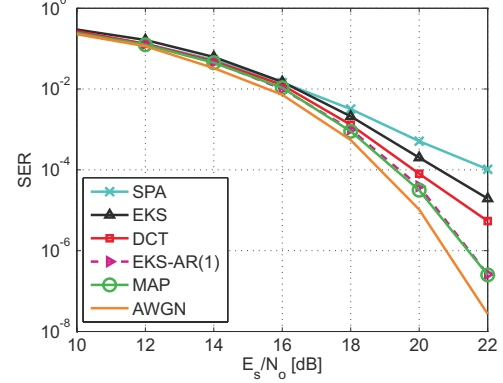


Fig. 4: SER comparison of different PN estimation methods after three estimation-detection iterations. Pilot density 21% and  $R_\zeta(0) = 10^{-3}[\text{rad}^2]$ .

and hence,  $\Sigma \approx -\tilde{\mathbf{H}}^{-1}(\theta^\dagger)$ . Employing the data-aided BCRB for estimation of colored PN derived in [7], and using the system model (3), it is straightforward to find the soft-input BCRB, and show that it is identical to the covariance of estimation error  $\Sigma$ .

#### REFERENCES

- [1] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 1748–1757, Sep. 2005.
- [2] H. Mehrpouyan, A. A. Nasir, S. D. Blostein, T. Eriksson, G. K. Karagiannis, and T. Svensson, "Joint estimation of channel and oscillator phase noise in MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sep. 2012.
- [3] P. Amblard, J. Brossier, and E. Moisan, "Phase tracking: what do we gain from optimality? particle filtering versus phase-locked loops," *Elsevier Signal Process.*, vol. 83, no. 1, pp. 151–167, Mar. 2003.
- [4] J. Bhatti and M. Moeneclaey, "Feedforward data-aided phase noise estimation from a DCT basis expansion," *EURASIP J. Wirel. Commun. Netw.*, Jan. 2009.
- [5] A. Demir, "Computing timing jitter from phase noise spectra for oscillators and phase-locked loops with white and  $1/f$  noise," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 9, pp. 1869–1884, Sep. 2006.
- [6] M.R. Khanzadi, A. Panahi, D. Kuylenstierna, and T. Eriksson, "A model-based analysis of phase jitter in RF oscillators," in *IEEE Intl. Frequency Control Symp. (IFCS)*, May 2012, pp. 1–4.
- [7] M.R. Khanzadi, D. Kuylenstierna, A. Panahi, T. Eriksson, and H. Zirath, "Calculation of the performance of communication systems from measured oscillator phase noise," *Accepted for publication in IEEE Trans. Circuits Syst. I, Reg. Papers*, Aug. 2013.
- [8] S. Song, A. C. Singer, and K. M. Sung, "Soft input channel estimation for turbo equalization," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2885–2894, Oct. 2004.
- [9] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*. Prentice Hall, Signal Processing Series, 1993.